

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



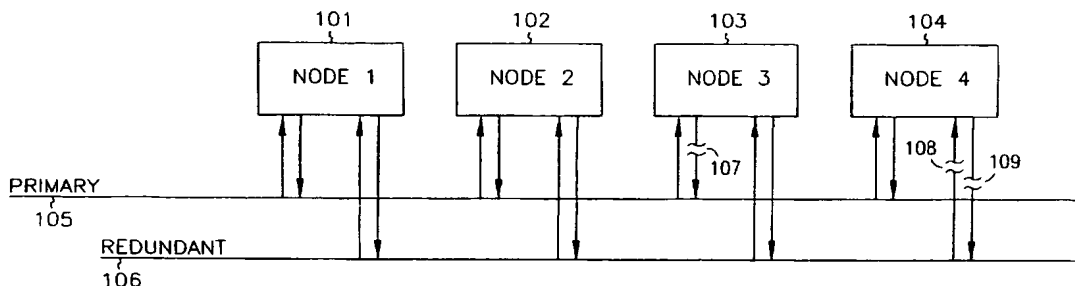
(43) International Publication Date
30 August 2001 (30.08.2001)

PCT

(10) International Publication Number
WO 01/63850 A1

- (51) International Patent Classification⁷: **H04L 12/40**, 29/14
- (21) International Application Number: PCT/US01/05834
- (22) International Filing Date: 22 February 2001 (22.02.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
09/513,010 25 February 2000 (25.02.2000) US
- (71) Applicant: **HONEYWELL INTERNATIONAL INC.**
[US/US]; 101 Columbia Road, P.O. Box 2245, Morristown, NJ 07960 (US).
- (72) Inventors: **HUANG, Jiandong**; 4365 Juneau Lane, Plymouth, MN 55446 (US). **GUSTIN, Jay, W.**; 6226 E. Joan De Arc Avenue, Scottsdale, AZ 85254 (US). **FREIMARK, Ronald, J.**; 13014 N. 62nd Street, Scottsdale, AZ 85254 (US). **KOZLIK, Tony, J.**; 2936 W. Morrow Drive, Phoenix, AZ 85027 (US). **SONG, Sejun**; Apartment F, 1026 27th Avenue S.E., Minneapolis, MN 55414 (US).
- (74) Agents: **CRISS, Roger, H.** et al.; Honeywell International Inc., 101 Columbia Avenue, P.O. Box 2245, Morristown, NJ 07960 (US).
- (81) Designated States (*national*): AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:**
— with international search report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: MULTIPLE NETWORK FAULT TOLERANCE VIA REDUNDANT NETWORK CONTROL



(57) Abstract: The present invention provides a method and apparatus for detecting and managing the state of a computer network comprising network nodes with redundant network connections, and for recovering from multiple network faults. In one embodiment, a network status table is employed in each node to manage data related to the network state between the node and other nodes in the network. In various embodiments, rerouting of data is managed independently such that a communication path is independently selected for sending data from a node to a connected node and for receiving data from the connected node. The invention in some embodiments is operable to route data through one or more intermediate nodes where direct connection between a pair of nodes is not possible.

WO 01/63850 A1

Multiple Network Fault Tolerance via Redundant Network Control

Field of the Invention

The invention relates generally to computer networks, and more specifically to a method and apparatus providing a fault-tolerant network having a redundant connection to network nodes able to detect and recover from multiple network faults.

Notice of Copending Applications

This application is related to the following copending applications, which are hereby incorporated by reference:

“Fault Tolerant Networking”, serial number 09/188,976; and

Atty. docket number 256.045us1

Background of the Invention

Computer networks have become increasingly important to communication and productivity in environments where computers are utilized for work.

Electronic mail has in many situations replaced paper mail and faxes as a means of distribution of information, and the availability of vast amounts of information on the Internet has become an invaluable resource both for many work-related and personal tasks. The ability to exchange data over computer networks also enables sharing of computer resources such as printers in a work environment, and enables

centralized network-based management of the networked computers.

For example, an office worker's personal computer may run software that is installed and updated automatically via a network, and that generates data that is printed to a networked printer shared by people in several different offices. The network may be used to inventory the software and hardware installed in each personal computer, greatly simplifying the task of inventory management. Also, the software and hardware configuration of each computer may be managed via the network, making the task of user support easier in a networked environment.

Networked computers also typically are connected to one or more network servers that provide data and resources to the networked computers. For example, a server may store a number of software applications that can be executed by the networked computers, or may store a database of data that can be accessed and utilized by the networked computers. The network servers typically also manage access to certain networked devices such as printers, which can be utilized by any of the networked computers. Also, a server may facilitate exchange of data such as e-mail or other similar services between the networked computers.

Connection from the local network to a larger network such as the Internet can provide greater ability to exchange data, such as by providing Internet e-mail access or access to the World Wide Web. These data connections make conducting business via the Internet practical, and have contributed to the growth in development and use of computer networks. Internet servers that provide data and serve functions such as e-commerce, streaming audio or video, e-mail, or provide other content rely on the operation of local networks as well as the Internet to provide a path between such data servers and client computer systems.

But like other electronic systems, networks are subject to failures.

Misconfiguration, broken wires, failed electronic components, and a number of other factors can cause a computer network connection to fail, leading to possible inoperability of the computer network. Such failures can be minimized in critical networking environments such as process control, medical, or other critical applications by utilization of backup or redundant network components. One example is use of a second network connection linking critical network nodes providing the same function as the first network connection. But, management of the network connections to facilitate operation in the event of a network failure can be a difficult task, and is itself subject to the ability of a network system or user to properly detect and compensate for the network fault. Furthermore, when both a primary and redundant network develop faults, exclusive use of either network will not provide full network operability. What is needed is a method and apparatus to detect and manage the state of a network of computers utilizing redundant communication channels.

Summary of the Invention

The present invention provides a method and apparatus for detecting and managing the state of a computer network comprising network nodes with redundant network connections, and for recovering from multiple network faults. In one embodiment, a network status table is employed in each node to manage data related to the network state between the node and other nodes in the network. In various embodiments, rerouting of data is managed independently such that a communication path is independently selected for sending data from a node to a

connected node and for receiving data from the connected node. The invention in some embodiments is operable to route data through one or more intermediate nodes where direct connection between a pair of nodes is not possible.

Brief Description of the Figures

Figure 1 shows a diagram of a computer network with multiple nodes having primary and redundant network connections, consistent with an embodiment of the present invention.

Figure 2 shows an example of a network status table, consistent with an embodiment of the present invention.

Figure 3 shows a flowchart of a method of managing the state of a network of nodes having primary and redundant network connections, consistent with an embodiment of the present invention.

Detailed Description

In the following detailed description of sample embodiments of the invention, reference is made to the accompanying drawings which form a part hereof, and in which is shown by way of illustration specific sample embodiments in which the invention may be practiced. These embodiments are described in sufficient detail to enable those skilled in the art to practice the invention, and it is to be understood that other embodiments may be utilized and that logical, mechanical, electrical, and other changes may be made without departing from the spirit or scope of the present invention. The following detailed description is, therefore, not to be taken in a limiting sense, and the scope of the invention is

defined only by the appended claims.

The present invention provides a method and an apparatus for detecting and managing the state of network connections to facilitate operation of a redundant network in the event of a network failure. The invention is capable of compensating for multiple network faults, including faults in both the primary and the redundant network. In some embodiments, the invention selects either the primary or the redundant network connection for communicating data between each pair of network nodes, such that the network may continue to be fully operational so long as at least one connection is operable to transmit data and one connection is operable to receive data between each pair of network nodes.

The invention in various forms is implemented using an existing network technology, such as Ethernet. In one such embodiment, two connections between each node are made via Ethernet connections — a primary network connection and a redundant network connection. In some such embodiments, off-the-shelf network adapters are utilized, and the invention controls the operation of the network adapters and manages communication via software executing on the computerized nodes. It is not critical for purposes of the invention which connection is the primary connection and which is the redundant connection, as the connections are physically and functionally similar. In the example embodiment discussed here, the primary and redundant network connections are interchangeable and are assigned names primarily for the purpose of distinguishing the networks from each other.

Figure 1 illustrates an exemplary network with four nodes 101, 102, 103 and 104. A primary network 105 and a redundant network 106 links each node to

the other nodes of the network, as indicated by the directional lines connecting the nodes to each of the networks. To understand how the invention is operable to compensate for multiple network failures, the connection from node 3 at 103 to primary network 105 is broken such that node 3 cannot transmit data to network 105 as shown at 107. Also, the connections linking node 4 at 104 to the redundant bus 106 are broken such that node 4 cannot receive data from the redundant bus as shown at 108 and cannot transmit data to the redundant bus as shown at 109.

In a typical redundant network system, failure of a single connection between the primary network and a node such as is shown at 107 would cause all nodes on the network to switch to communicating via the redundant bus 106. In the network configuration shown in Figure 1, connections between node 4 and the redundant bus are also inoperable, making operation of the network using the redundant bus impossible. Such multiple failures make the network inoperable when exclusively using either the primary or redundant bus.

The present invention provides a solution to this problem and enables communication between all network nodes during multiple failures such as are shown in Figure 1 by use of network status data and intelligent routing of data. In some embodiments of the invention, the network status data is stored in a network status table as shown in Figure 2.

Figure 2 illustrates an example of a network status table for node 3 of the network of Figure 1, and contains data indicating the ability of node 3 to receive data from other nodes and the ability of other nodes to receive data from node 3. Specifically, the "Received Data OK" columns indicate the ability of node 3 to receive data from each of nodes 1, 2 and 4 on both the primary and redundant

networks. The table indicates with an “X” that node 3 cannot receive data from node 4 over the redundant network connection, and indicates that node 3 can receive data from all other nodes via both the primary and redundant network connections with an “OK”. The “X” indicating node 3's inability to receive data from node 4 is the result of the broken data transmit connection 109 between the redundant network 106 and node 4 (104).

The “Other Node Report Data” columns represent the data reported to node 3 by other nodes regarding the ability of the various other nodes to receive data from node 3. Because node 3's connection to the primary network 105 is broken at 107 such that node 3 cannot send data over the connection, nodes 1, 2 and 4 are unable to receive data from node 3 on the primary network and so an “X” indicates a node 3 failure for each of these nodes. Also, the data connection between node 4 and the redundant network is broken at 108 such that node 4 cannot receive data from the redundant network, so an “X” also indicates that node 4 is unable to receive data from node 3 in the node “4” column of the “Node 3 Redundant” row.

The determination of whether a node can receive data from another node is made in various embodiments using special-purpose diagnostic data signals, using network protocol signals, or using any other suitable type of data sent between nodes. The data each node provides to other nodes to populate the “Other Node Report Data” must necessarily be data which includes the data to be communicated between nodes, and is in one embodiment a special-purpose diagnostic data signal comprising the node data to be reported.

From the data in the network status table of Figure 2, the state of the various network connections can be determined and a suitable connection for

communication between each pair of network nodes can be selected. In the example of Figures 1 and 2, nodes 1 and 2 are fully operational and may use either connection to communicate, and nodes 3 and 4 each have a fully operational connection to either the primary or redundant networks. Therefore, only nodes 3 and 4 are unable to communicate over either the primary or redundant network exclusively. Node 3 cannot send data to the primary network, and node 4 cannot send or receive data from the redundant network, but node 3 can receive data from node 4 via the primary network. In some embodiments of the invention, node 3 cannot send data to node 4 because no operable direct path over either the primary or redundant networks exists to send data.

In other embodiments of the invention, node 3 may transmit the data to node 4 via another node with an "OK" indication for either network in the "Other Node Report Data" rows of the table such as node 1 or node 2. In such embodiments, the "OK" nodes or intermediate nodes are known to be able to receive data from node 3, and can retransmit the data to node 4 via their fully functional primary network connections. This allows communication between two nodes where multiple network failures prevent direct communication between two nodes. In further embodiments, the intermediate node to which the data is routed is selected via polling the intermediate nodes to select a node that indicates it is able to retransmit data to node 4 by evaluation of the data in each of the intermediate nodes' network status table. In various embodiments of the invention, the intermediate nodes may comprise networked computers as in the example above, may comprise a direct connection between networks, may comprise a router or bridge, may comprise a special-purpose intermediate node hardware device, or may

be implemented in any other way that provides the ability to suitably communicate signals between the two networks.

Figure 3 is a flowchart illustrating a method of practicing one embodiment of the present invention. At 301, each node determines the state of the primary network connection linking it to each other node. Also, the state of the redundant network connection linking each node to each other node is determined at 302. The state of the primary and redundant connections between each pair of nodes can be determined in various embodiments by searching the connections for existing data such as valid data or protocol packets, or by use of special-purpose diagnostic messages. This network connection state data is used at 303 to build the "Received Data OK" portion of a network status table for each node, and the nodes exchange data with each other at 304 to complete the "Other Node Report Data" portion of the network status table. The network status table is updated regularly, and is monitored at 305 to determine whether a network connection has failed and requires rerouting of data.

At 306, the node determines by examination of the network status table whether a direct connection for transmitting and receiving data between the pair of nodes with a failed connection can be made. If a connection can be made, such as by transmitting data via the primary network connection and receiving data through the redundant network connection, the data is rerouted through the direct connections at 307 and monitoring for additional failures resumes at 305. If a direct connection cannot be made, data is rerouted through one or more intermediate nodes at 308 to facilitate communication, as was described in accordance with the multiple network failure example illustrated in Figures 1 and

2. Again, once a data path through one or more intermediate nodes has been selected monitoring for additional network failures resumes at 305.

The present invention provides a method and apparatus that enable a network with primary and redundant network connections to manage routing of data through the network such that multiple network failures can be compensated for. In some embodiments, the invention includes rerouting data that cannot be transferred directly between two nodes to intermediate nodes which are able to facilitate communication between the nodes. The invention also incorporates construction and use of a network status table in some embodiments for managing data related to the network state. The invention includes in various embodiments a method for managing the state of the network, software for execution on a computer for managing the state of the network, and a hardware network interface that is operable to manage the state of the network.

Although specific embodiments have been illustrated and described herein, it will be appreciated by those of ordinary skill in the art that any arrangement which is calculated to achieve the same purpose may be substituted for the specific embodiments shown. This application is intended to cover any adaptations or variations of the invention. It is intended that this invention be limited only by the claims, and the full scope of equivalents thereof.

Claims

1. A method of managing the state of a computer network with redundant network connections, comprising:

determining the state of a primary network connection between each pair of networked nodes;

determining the state of a redundant network connection between each pair of networked nodes; and

selecting either the primary network connection or the redundant network connection for sending and receiving data between each pair of networked nodes, such that the network path selected to be used to communicate is selected independently based on the determined network states for each pair of networked nodes.

2. The method of claim 1, further comprising building a network status table that indicates results of determining the state of primary and redundant network connections between each pair of networked nodes.

3. The method of claim 2, wherein the network status table comprises data representing network status based on data received at a node from other network nodes.

4. The method of claim 3, wherein the data received at a node from other

networked nodes comprises a diagnostic message.

5. The method of claim 4, wherein the data received at a node from other networked nodes comprises data representing the ability of the other nodes to receive data from other different network nodes.

6. The method of claim 2, wherein the network status table comprises data representing network status based on a node's ability to send data to other nodes.

7. The method of claim 3, wherein the network status table further comprises data representing network status based on a node's ability to send data to other nodes.

8. The method of claim 1, wherein selecting the primary or redundant network connection for communication between each pair of networked nodes comprises:

- selecting the primary network connection if the state of the primary network connection is determined to be operable; and
- selecting the redundant network connection if the state of the primary network connection is determined to be inoperable.

9. The method of claim 1, wherein selecting the primary or redundant network connection for communication between each pair of networked nodes comprises:

- selecting the primary network connection to transmit data if the state of the primary network connection is determined to be operable to transmit data;
- selecting the primary network connection to receive data if the state of the

primary network connection is determined to be operable to receive data;

selecting the redundant network connection to transmit data if the state of the primary network connection is determined to be inoperable to transmit data;

and

selecting the redundant network connection to receive data if the state of the primary network connection is determined to be inoperable to receive data.

10. The method of claim 1, wherein selecting a connection for sending and receiving data between each pair of network nodes comprises selecting a connection for sending and receiving data from a first node to one or more connected intermediate nodes, and selecting a connection for sending and receiving data from an intermediate node to a second node.

11. A computer network interface, the interface operable to:

determine the state of a primary network connection between the network interface and the network interfaces of other network nodes;

determine the state of a redundant network connection between the network interface and the network interfaces of other network nodes; and

select either the primary network connection or the redundant network connection for communication with each of the other network nodes, such that the network connection selected is selected independently based on the determined network states for each other network node.

12. The computer network interface of claim 11, the interface further comprising a

network status table that indicates results of the determination of the state of the primary and redundant network connections between the computer network interface and the network interfaces of other network nodes.

13. The computer network interface of claim 12, wherein the network status table comprises data representing network status based on data received at a node from other network nodes.

14. The computer network interface of claim 13, wherein the data received at a node from other networked nodes comprises a diagnostic message.

15. The computer network interface of claim 14, wherein the data received at a node from other networked nodes further comprises data representing the ability of the other nodes to receive data from other different network nodes.

16. The computer network interface of claim 12, wherein the network status table comprises data representing network status based on a node's ability to send data to other nodes.

17. The computer network interface of claim 13, wherein the network status table further comprises data representing network status based on a node's ability to send data to other nodes.

18. The computer network interface of claim 11, wherein selecting either the

primary network connection or the redundant network connection for communication with each of the other network nodes comprises:

selecting the primary network connection if the state of the primary network connection is determined to be operable; and

selecting the redundant network connection if the state of the primary network connection is determined to be inoperable.

19. The computer network interface of claim 11, wherein selecting either the primary network connection or the redundant network connection for communication with each of the other network nodes comprises:

selecting the primary network connection to transmit data if the state of the primary network connection is determined to be operable to transmit data;

selecting the primary network connection to receive data if the state of the primary network connection is determined to be operable to receive data;

selecting the redundant network connection to transmit data if the state of the primary network connection is determined to be inoperable to transmit data; and

selecting the redundant network connection to receive data if the state of the primary network connection is determined to be inoperable to receive data.

20. The computer network interface of claim 11, wherein selecting a connection for sending and receiving data between each pair of network nodes comprises selecting a connection for sending and receiving data from a first node to one or more connected intermediate nodes, and selecting a connection for sending and

receiving data from an intermediate node to a second node.

21. A machine-readable medium with instructions thereon, the instructions when executed on a computer operable to cause the computer to:

determine the state of a primary network connection between the network interface and the network interfaces of other network nodes;

determine the state of a redundant network connection between the network interface and the network interfaces of other network nodes; and

select either the primary network connection or the redundant network connection for communication with each of the other network nodes, such that the network connection selected is selected independently based on the determined network states for each other network node.

22. The machine-readable medium of claim 21, the instructions further operable to cause a computer to create and maintain a network status table that indicates results of the determination of the state of the primary and redundant network connections between the computer network interface and the network interfaces of other network nodes.

23. The machine-readable medium of claim 22, wherein the created network status table comprises data representing network status based on data received at a node from other network nodes.

24. The machine-readable medium of claim 23, wherein the data received at a

node from other networked nodes comprises a diagnostic message.

25. The machine-readable medium of claim 24, wherein the data received at a node from other networked nodes further comprises data representing the ability of the other nodes to receive data from other different network nodes.

26. The machine-readable medium of claim 22, wherein the created network status table comprises data representing network status based on a node's ability to send data to other nodes.

27. The machine-readable medium of claim 23, wherein the network status table further comprises data representing network status based on a node's ability to send data to other nodes.

28. The machine-readable medium of claim 21, wherein selecting either the primary network connection or the redundant network connection for communication with each of the other network nodes comprises:

selecting the primary network connection if the state of the primary network connection is determined to be operable; and

selecting the redundant network connection if the state of the primary network connection is determined to be inoperable.

29. The machine-readable medium of claim 21, wherein selecting either the primary network connection or the redundant network connection for

communication with each of the other network nodes comprises:

selecting the primary network connection to transmit data if the state of the primary network connection is determined to be operable to transmit data;

selecting the primary network connection to receive data if the state of the primary network connection is determined to be operable to receive data;

selecting the redundant network connection to transmit data if the state of the primary network connection is determined to be inoperable to transmit data;
and

selecting the redundant network connection to receive data if the state of the primary network connection is determined to be inoperable to receive data.

30. The machine-readable medium of claim 21, wherein selecting a connection for sending and receiving data between each pair of network nodes comprises selecting a connection for sending and receiving data from a first node to one or more connected intermediate nodes, and selecting a connection for sending and receiving data from an intermediate node to a second node.

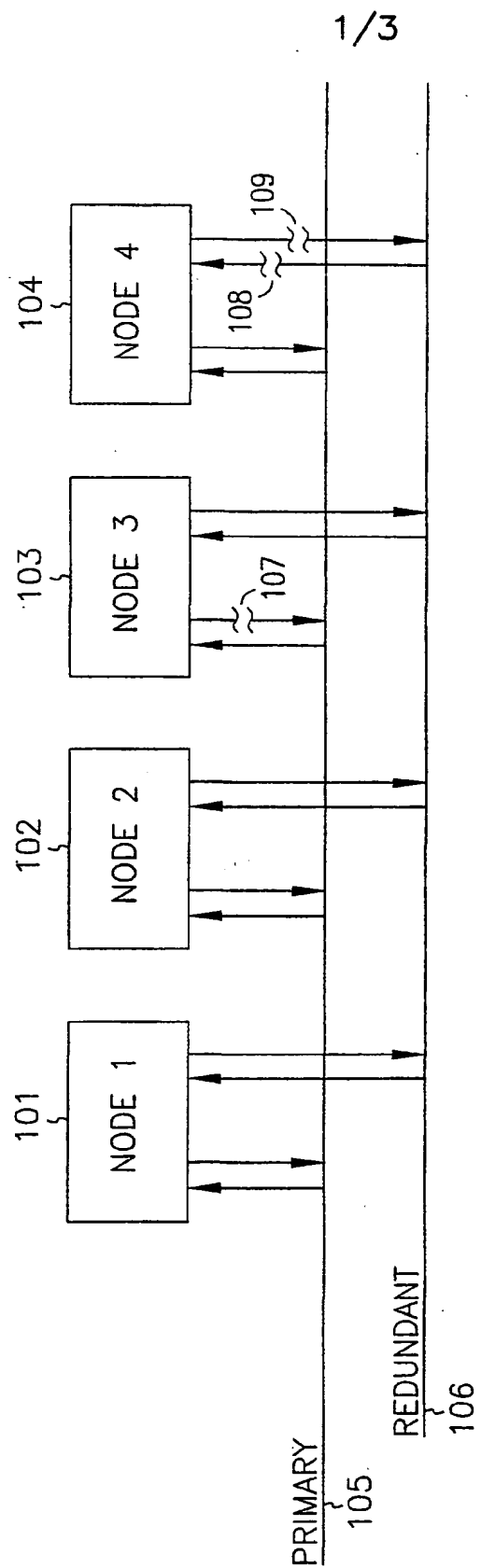


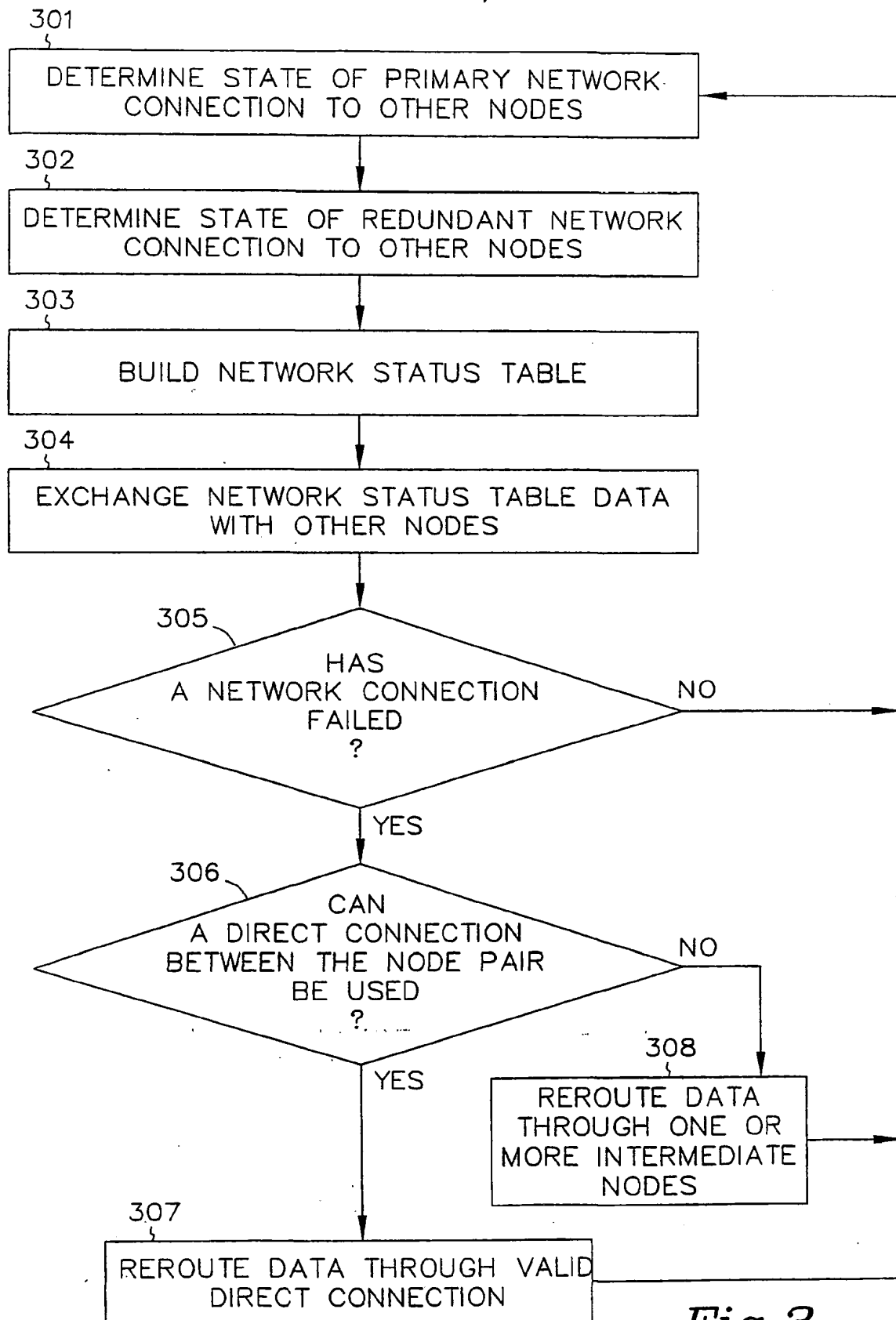
Fig.1

NODE 3 NETWORK STATUS TABLE

RECEIVED DATA OK			OTHER NODE REPORT DATA	
NODE#	PRIMARY	REDUNDANT	NODE 3 PRIMARY	NODE 3 REDUNDANT
1	OK	OK	X	OK
2	OK	OK	X	OK
-	-	-	-	-
4	OK	X	X	X

Fig.2

3/3

*Fig.3*

INTERNATIONAL SEARCH REPORT

International: ation No
PCT/US 01/05834A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 H04L12/40 H04L29/14

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, COMPENDEX, IBM-TDB, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 96 08098 A (ELIN ENERGIEANWENDUNG) 14 March 1996 (1996-03-14) abstract; figure 1 page 4, line 20 -page 6, line 5 -----	1-30
A	US 5 684 807 A (BIANCHINI JR RONALD P ET AL) 4 November 1997 (1997-11-04) page 2, line 24-35 column 5, line 61 -column 6, line 16; figure TABLE1 -----	2, 4

☐ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- * & * document member of the same patent family

Date of the actual completion of the international search

1 June 2001

Date of mailing of the international search report

11/06/2001

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Hardelin, T

INTERNATIONAL SEARCH REPORT

Internat cation No
PCT/US 01/05834

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 9608098	A	14-03-1996	AT	983 U	26-08-1996
US 5684807	A	04-11-1997	US	5325518 A	28-06-1994
			AU	1756592 A	02-11-1992
			EP	0578730 A	19-01-1994
			MX	9201501 A	01-10-1992
			WO	9217961 A	15-10-1992
			ZA	9202378 A	24-02-1993